DOCUMENT RESUME

ED 229 405 TM 830 301

AUTHOR Dockrell, W. B.
TITLE Applicability of Standards for Evaluations of Educational Programs, Projects and Materials in an International Setting: Qualitative Research.
PUB DATE Apr 83
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, April 11-15, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Data Analysis; Foreign Countries; *International Studies; Privacy; *Program Evaluation; Research Methodology; Research Needs; *Research Problems; *Standards
IDENTIFIERS Collusion; Conflict of Interest; *Qualitative Research; Scotland (Edinburgh); *Standards for Evaluation Educ Prog Proj Materials

ABSTRACT

This paper is devoted to an analysis of the "Standards for Evaluation of Educational Programs, Projects, and Materials" as they apply to qualitative data. The most difficult ethical problem arises from the conflict between the right to privacy of research subjects and the right to know of audiences. There are two technical points in the standards document which should be amplified. The first concerns the verification of data: the need to specify procedures to ensure credibility for qualitative studies. The second concern is the specification of quantitative techniques for the analysis of qualitative data: the need to define more clearly the steps that should be taken to ensure that qualitative data can be tested rigorously. Finally there is the problem of collusion, the danger that the qualitative evaluator will select from his material the data which is in conformity with his own predilictions or presuppositions, the expectations of his audience, or the requirements of his sponsors. This problem may be acute in some parts of the world. To be of maximum international value the standards report needs to be extended and to pay much greater attention to problems arising with qualitative evaluation. (Author/PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED229405

APPLICABILITY OF STANDARDS FOR EVALUATIONS OF EDUCATIONAL PROGRAMS, PROJECTS AND MATERIALS IN AN INTERNATIONAL SETTING: QUALITATIVE RESEARCH

By W B Dockrell
Scottish Council for Research in Education
Edinburgh

Contents

Page

Summary 1
1. The Prevalance of Qualitative Research 2
2. Conflicts of Interest between Subjects and Audiences 6
3. Verifying Qualitative Information 11
4. Quantitative Analyses of Qualtitative Data 12
5. The Risk of Collusion 14
6. Conclusion 15
7. References 16

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

W. B. Dockrell

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 830 301

SUMMARY

Qualitative studies are much more the practice in Europe and increasingly elsewhere than in the United States. For an international audience therefore there needs to be much greater emphasis on qualitative techniques than there is in the standards report.

The most difficult ethical problem arises from the conflict between the right to privacy of research subjects and the right to know of audiences. By its nature qualititative research focuses on the particular, the detailed analysis of the specific incident. It is therefore often difficult to conceal the identity of the individual organisations and people concerned. Even a confidential report might work to the detriment of the subjects studied.

Published reports pose the same problem. Have subjects the right to have material they believe to their disadvantage excised from a report even if that renders the report valueless?

There are two technical points in the standards document which should be amplified. The first concerns the verification of data. We need to specify procedures to ensure credibility for qualitative studies.

The second concern is the specification of quantitative techniques for the analysis of qualitative data. The statement in the standards document is too weak. We need to define much more clearly the steps that should be taken to ensure that qualitative data can be tested rigorously.

Finally there is the problem of collusion, the danger that the qualitative evaluator will select from his material the data which is in conformity with his own predilictions or presuppositions, the expectations of his audience or the requirements of his sponsors. This problem may be acute in some parts of the world.

To be of maximum international value the standards report needs to be extented and to pay much greater attention to problems arising with qualitative evaluation.

1. The Prevalence of Qualitative Research

This paper is devoted to an analysis of the standards as they apply to qualitative data. Why, you may ask, should it be necessary to refer to qualitative data in an international symposium on standards? Most of the general issues covered in the book apply as much to qualitative information as to quantitative; there is a section devoted specifically to the analysis of qualitative information; and surely there is qualitative research in the United States as elsewhere. True, but it seems to me that there are specific issues which arise in the case of qualitative research which are not adequately covered in the broader discussion. The section on qualitative research is brief - less than five pages - and that is because qualitative research, although it exists in the United States, does not have the importance there that it does elsewhere.

One of the most trenchant analyses of the limitations of both experimental and correlational studies is Cronbach's "Beyond the two Disciplines of Scientific Psychology". (1975). He concludes that article by arguing for the approach to knowledge of mankind which underlies the qualitative approach.

> "Though enduring systematic theories about man in society are not likely to be achieved, systematic enquiry can realistically hope to make two contributions. One reasonable aspiration is to assess local events accurately, to improve short run control ... The other reasonable aspiration is to develop explanatory concepts, concepts that will help people use their heads". (Page 126)

He goes on to assert:

> "The special task of the social scientist in each generation is to pin down the contemporary facts". (Page 126)

One major task of the evaluator is "to pin down the contemporary facts" and he may do it for either of the purposes that Cronbach lists, that is, "to improve short run control" and/or "to develop explanatory concepts". That is an important distinction to which we will be returning, for it seems to me that much evaluation is concerned with the former purpose and that purpose alone, that is, understanding of a specific situation without necessarily contributing to a broader understanding. In spite of Cronbach's eloquent plea for a different approach, he does not seem to have affected those

to whom his paper was addressed, that is, educational psychologists. The 1982 volume of the Journal of Educational Psychology has nearly one hundred articles of which only one can be said to take the position advocated by him.

Nor does Cronbach seem to have influenced American educational researchers. An analysis of the four numbers of the American Educational Research Journal published in 1982, shows that the articles contained in it are overwhelmingly reports of experimental or correlational studies. The standard article consists of an introduction, a review of previous evidence or a statement of the model or theory to be tested, a section on methods, one on results and finally, one containing conclusions or discussions. The pattern is so standard that it is hard to find one that deviates from it. There are though a few articles of two other types, one consisting of meta-analyses which themselves are empirical in that they consist of a structured analysis of previously published materials, and the other consists of methodological papers which analyse the merits of particular experimental or correlational techniques.

A comparison with the British Educational Research Journal is revealing; the British Journal has only two issues per year, but that is far from being the most striking difference between them. Most of the articles in the British Journal report experimental or correlational studies which could as easily have appeared in the American Journal. There are too, articles concerned with technique which could have appeared in the AERJ. There are, however, also articles written from a different stance like Hewit's (1982), "A Critique of Research Methods in the Study of Reading Comprehension" which might in principle have appeared in the Review of Educational Research, but is unlikely to have done so because it is not a meta-analysis in the sense of analysing quantitatively the European and American research, rather it is analytic arguing for "naturalistic research using observational techniques".

> "We need to describe what readers do while reading and after they have read, and to theorise about and discover what influences their comprehension and their conception of comprehension; we need to discover what kinds of meanings reading has for different individuals, and what the social and political determinants are of these meanings". (Page 9).

Perhaps an article like Chamber's (1982) "School Based Curriculum Research" might have found an American audience, though its emphasis on providing "an ideological basis for justifying school-based curriculum research" (Page 133) might well have seemed alien.

What is significant is not that there are studies of the kind that Cronbach advocates carried out in European countries, it is the extent to which the approach which he advocates permeates European educational research and evaluation. It is not merely the tradition of geisteswissenschaft and not only in Germany, but that this tradition has deeply influenced what educational researchers do at least in the course of the last ten years, as Kallos (1980) has pointed out. It is not the only, and in some countries, not the dominant form of evaluation, but it takes its place in evaluation practice with other traditions.

Studies of this kind are typical of a widespread approach to educational evaluation in Britain. They are supported financially by the major funders including central government, local education authorities and major foundations like the Social Science Research Council, Leverhulme and Nuffield. It is worth emphasising that while this is an approach to evaluation, it is also an approach to research. It involves a definition of what knowledge is in the social sciences (including education) and how it is likely to be increased.

The group which has practised this approach and which has been most active in analysing it has been associated with the University of East Anglia, and particularly with the late Lawrence Stenhouse and his colleagues, including David Hamilton.

The Scottish Council published a case-study of the transition to open-plan in a single primary school carried out by David Hamilton (1977) when he was on the Council staff. In his preface to that book, Hamilton harks back to an earlier publication by the Scottish Council in 1974. Hamilton comments that the earlier book

> "Did not set out to test a range of pre-specified hypotheses or even to survey every aspect of open-plan schooling. Instead it tried to respond in an accessible manner to some of the questions posed at that time by teachers and administrators". (Page IV)

He goes on to say that his own study

> "Utilises a similar research perspective. It identifies the general phenomena of open-plan schooling and locates them in the day to day work of one school ... Its principal concern is to contribute - in a concise but sensitising manner - to wider discussion about the problems and possibilities of open-plan schooling". (Page V)

In a more recent publication another member of the group, Elliot (1981), now at Cambridge University, reporting an evaluation of different approaches to school gives his reason for adopting the case-study approach.

> "First, we wanted our research to contribute to the development of self-accounting practices in particular schools ... If the project was to help schools in developing their responsiveness to local audiences, then it needed to understand the problems and effects of self-accounting in the light of the particular circumstances they faced.
>
> Second, since the self-accounting procedure adopted by schools, and the problems and effects which arise as a result will depend on their own situation, we were sceptical about the possibility of producing a definitive set of "research findings" in the form of generalisations which could be applied to all schools irrespective of the settings in which they operated.
>
> However, we felt that by comparing and contrasting the "problems and effects" of self-accounting across a small number of cases and explaining them in terms of similarities and differences between settings, we could generate some ideas and hypotheses which staff in schools outside the project could usefully explore and examine in relation to their circumstances". (Page XIV)

A third theme running through the European studies parallels Cronbach's third point, a distrust of theory. Cronbach says

> "The goal of our work I have argued here is not to amass generalisations, atop which a theoretical tower can some day be erected". (Page 126)

Macdonald and Walker (1976) make the same point in their Analysis of Curriculum Change with a warning

> "In some fields of education, theory is fully developed, sophisticated, complex and demanding. Such theory can trap the beginner, offering him elegant substitutes for his own analyses of and reflections upon the phenomena they address or it may undermine the student who seeks his own understanding of the field by posing a set of baffling obstacles between him and his data". (Page 4)

Not all European researchers take the position of the theorists (if they would let me call them that) that I have just quoted, but it is a wide-spread approach. Most of the studies currently being carried out by the Scottish Council are qualitative in their approach. Some of them are evaluations of current practice or new procedures commissioned by customers as a basis for making change. Perhaps the best example is a book with that title Making the Change (Cumming et al 1981) which is a study of guidance, discipline and sanctions in a number of schools which had eliminated corporal punishment.

It was funded jointly by the Scottish Education Department (Central Government) and the local education authorities. Corporal punishment has in the past been a much more common feature of Scottish than American schools, though now partly, perhaps largely as a result of this evaluation, it has already been banned by the two largest authorities in Scotland covering between them something like two-thirds of all schools, and the same ban will be instituted in the rest of the country shortly.

The study, in accord with the principles in the standards report, used quantitativ as well as qualitative techniques. There were systematic classroom observations; questionnaires and structured interviews as well as informed observations and unstructured interviews. There were, in the terms of the standards report

> "facts and interpretations that are in narrative rather than numerical in form (unobtrusive data) gathered intentionally, or ... quite unexpectedly". (Page 130).

In Cronbach's words there was

> "Intensive local observation (which) goes beyond discipline to an open-eyed, open-minded appreciation of the surprises nature deposits in the investigative net". (Page 125).

Not only is this the pattern in Europe, but it is one which is being adopted elsewhere. A consultant to the World Bank, for example, argued that qualitative studies were the appropriate way of evaluating a major educational reform in a developing country and that this approach should be the pattern for future evaluations rather than more traditional quantitative approaches. In most countries qualitative studies need to be the focus of a book on standards and not an appendix.

There are serious problems for the application of the standards to the kind of evaluations described. Some of them apply to all qualitative research, some apply to case study research in particular.

2. Conflicts of Interest between Subjects and Audiences

My first comment relates to the first section of the report (A1) audience identification and relates it to later sections, the public's right to know and the right of human subjects (C4 and 5). How can we reconcile the rights to know of a customer commissioning a study and other interested parties and the right of those involved in the study to confidentiality?

Many of our studies are single case studies of the kind described by Cronbach, 8

short run control".(Page 126) We have recently completed or are in the process of completing, two studies of individual colleges at the request of the colleges themselves of a programme in a single local education authority at the request of the authority, and an experimental programme in a single school at the request of the school and the foundation funding the programme. That is a substantial proportion of our work and that of some other comparable evaluation organisations. For example, the evaluations of the European Economic Community action programme entitled Transition from School to Working Life (cf. the report on the Clydebank EEC project Stronach and Weir 1983) and the UK wide educational priority areas evaluation which was probably the largest evaluation of education funded within the UK. (Halsey 1972) All these cases involve studies of identifiable single institutions and identifiable if not identified individual people.

As publishers we issued a report which was a study of a single school carried out by an evaluator working for another organisation. (Gordon 1981) The school was not named and was described in a way that was meant to conceal its identity. Teachers were observed in their teaching. Again they were referred to in a way which was meant to conceal their identity, but a sequence of teaching was described and commented on critically. Perhaps a critical comment was inappropriate but even a simple description would be sufficient to have raised questions in many minds. The teacher concerned complained that only he had the opportunity to teach that particular sequence and therefore was identifiable to his colleagues and to the administrators within the school and that the report had prejudiced his professional future. We had not thought that he was identifiable. He thought he was. It seemed to the author of the report that this was a particularly significant piece of information which was relevant to a wider constituency. Had the individual concerned the right to insist that it was excised?

We were recently commissioned to make an evaluation of a new experimental programme of professional training in one college. There were two audiences particularly concerned with this experiment, the college itself which would wish to review and possibly revise its traditional practice in the light of the findings of the action research. The other body was the Government Department with responsibility for overseeing the training, certification and indeed to a large extent the employment of the professional group concerned, for the successful outcome of the experiment would have implications for the practice in other colleges and universities in the country.

The evaluation was a case study, qualitative in nature, based on interviews with students, in both the experimental and the traditional form of programme, with the staff of the college and with the professionals involved in the organisation and supervision of the students' field training experiences.

The conclusions are based on an interpretation of the interview data. The students and the professionals involved in the field practice were a sufficiently large group for their comments not to be attributable to any particular individual, but a senior administrator of the college who made some trenchant comments, the head and members of the college department concerned and the principal researcher were all easily identifiable.

The experiment was highly controversial. The reputation and the professional future of the researcher was to a large extent tied up with this particular study. The college department felt itself threatened. At one point the evaluation described the members of the department as fighting for their professional lives. The administration of the college were in a position where they had to make comments on the adequacy of current practice which was inevitably seen to be critical of existing members of staff. All of this of course was highly relevant to the college and the government department. It was possible to prepare a private report which although it omitted names was perfectly clear as to the source and therefore the validity of the comments which it quoted and the implications for specific practices. There would be no doubt in the readers minds who the people were and what the practices were.

Presumably those involved in the experiment had realised that they were putting their professional necks on the block and accepted that the evaluation would have implications for them. They clearly have a right to read and react to the document and to see that their responses are incorporated in the final report to the sponsors but have they a right to insist that certain information about themselves which they feel to be to their disadvantage should be struck from the report. Have they a right to claim in the words of the standards "privacy of certain thoughts and information, confidentiality of some information" (Page 81), even if by doing so they should vitiate the whole report.

If not, then there is no problem. The report can go forward. If they have at what stage have they this right? As the report points out in its section on the analysis of qualitative information because of the nature of qualitative information "the evaluator must assure the accuracy of his findings

... by subjecting inferences to independent verification", (Page 131) but who in circumstances like this can provide independent verification without getting to know the individuals concerned?

We have for our projects, indeed we are required to have by our major funder, an advisory committee which among other things reads the drafts of the interim and final reports. In this case the committee consisted of representatives of the government department, of the professionals in the field and of other colleges. They have seen and hold copies of the unexpurgated report. We cannot wipe the tapes clean.

If the subjects of the study have no objection a report can be submitted to the college and to the government department but what about the others with a right to know. What about the profession? Others involved in training members of the profession? The students? Indeed the public who benefit from the activities of adequately or inadequately trained professionals? It is our practice and indeed we have a formal requirement, that the findings of our researches be made available in some form. What can be said in a public report that will not be clearly related to a particular college, to a particular department, to a particular experiment and to the people concerned, possibly to their detriment and disadvantage? How do we deal with these issues?

I am not suggesting that I have a solution, merely that I am aware of the problem and that this is a major problem for those who are involved in this kind of evaluation.

Problems arise not only with evaluations based on a single instance as in the example given in the standards document and as is common practice. It can be based on a series of "intensive local observations" (Cronbach, Page 125). Here the danger is less acute. It is possible to report the findings by categories across cases - sometimes. There is always the unique combination of circumstances that can refer only to one particular case and which a late 20th century Sherlock Holmes can identify. Does it matter? Well it does if it does, that is, if it is seen by the organisation or the person concerned to be working to their disadvantage. If not, of course it does not matter.

A recent report that we published concerned the use or rather the non-use of corporal punishment in a small number of secondary schools (Cummings et al 1980). Of all the secondary schools that we could find in Scotland at that time which did not have corporal punishment, only two were city schools serving neighbourhoods

where the people were deprived, dis-advantaged, depressed or, as we say in Britain, poor. Both of these had distinctive characteristics which made it difficult to disguise them. Both were experimental schools, well-known for their new programmes. It would have been impossible to hide from the informed reader which school it was that was being referred to.

This research falls into Cronbach's second category, that is, research which is designed "to develop explanatory concepts that will help people use their heads" (Page 126). Its audience therefore was not the institutions in which the data was gathered but rather other institutions where practices were different.

If I might quote Cronbach again this researcher was in the position described as follows:-

> "An observer collecting data in one particular situation is in a position to appraise a practice or proposition in that setting, observing effects in context. In trying to describe and account for what happened he will give attention to whatever variables were controlled but he will also give equally careful attention to uncontrolled conditions, to personal characteristics and to events that occurred during treatment and measurement. As he goes from situation to situation the first task is to describe and interpret the effect anew in each locale perhaps taking into account factors unique to that locale of a series of events ... as results accumulate a person who seeks understanding will do his best to trace how the uncontrolled factors could have caused local departure from the modal effect, that is, generalisation comes late and the exception is taken as seriously as the rule" (Page 124 - 125).

Now if the report is to be so written that it "will help people use their heads" (Page 126) then it presumably will have to contain information which refers to "uncontrolled conditions, to personal characteristics" and to "factors unique to that locale". It is impossible in these case studies to disguise from an informed audience, (that is, the audience to which the report was addressed) which were the schools involved, what were the practices in those schools and indeed what were the "local departures from the modal effect". The reader therefore would have an opportunity to make his own judgement about the schools and particularly about the administrators concerned. Should those involved have the right to require the removal of any material about themselves to which they object, even if this were to weaken the report to the point of rendering it valueless? What is the balance between the rights of the subjects in the study and of the educational community for whose benefit the study was carried out?

The problem is to balance the competing claims of different principles - the principles of the right to know and the principle of confidentiality.

3. Verifying Qualitative Information

Let me now turn to an entirely different issue. That is, the problem of verification. It is one that is referred to in several parts of the section on qualitative information. A research psychologist commented to me recently on reading a report of a qualitative study that it was all very convincing but how did she question the data? How did she know whether it was an accurate representation of what went on, whether the quotations given were a fair sample of the comments made during the interviews and so on. If I may refer to another source, in the film, Gandhi, the Mahatma is asked by a reporter how he should handle his report of a particular incident. The reply was "I want you to document coldly and rationally what is happening". That is part of the duty of a qualitative evaluator. The problem for the reader is to know whether the report has documented "coldly and rationally" or whether it has distorted unconsciously to fit pre-suppositions or indeed consciously for political or other ends.

The standards document states emphatically that "the evaluator must assure the accuracy of findings by seeking confirmatory evidence from more than one source and subjecting influences to independent verification". (Page 131). Indeed, but easier said than done. How does one get the confirmatory evidence of, for example, the behaviour of one teacher, with one class on one occasion (to refer back to a previous example). It is impossible in the case of qualitative classroom observation to present the evidence to another researcher especially if the information "came to one's awareness quite unexpectedly".(Page 130). It may be crucial information but the circumstances in which it occurred may never be repeated. The data are often, by their very nature, not repeatable, not confirmable, subject to misinterpretation and to misrepresentation. We need much clearer guidelines.

I do not know what those guidelines might be but I do know that until we can be clearer there is a problem with evaluator credibility. The standards say that "persons conducting the evaluation should be both trustworthy and competent to perform the evaluation". (Page 24). There would be a continuing and perhaps increasing tendency to accept qualitative evaluations from those who one believes to be "both trustworthy and competent" and to be sceptical about the reports of

evaluators one does not know or about whom one has doubts.

4. Quantitative Analyses of Qualitative Data

There is though one area where much firmer guidelines can be laid down. We now have enough experience to say that unless certain procedures have been followed we should be at best sceptical. That is in the analysis of qualitative information. The document talks about categories and testing the categories for validity and reliability. We can and in a statement of guidelines it seems to me we should say that there are procedures which should be followed.

Let me give you an example. In one study that we are now conducting we are analysing interviews with a number of categories of schools and education office staff. There is a team of three, two field workers who carry out the interviews and a principal investigator. Each interviewer analyses his own materials suggesting tentative categories and applying these to a sample of the transcripts. The field workers then analyse a sample of each other's transcripts, applying and revising their own set of categories. The transcripts which have been analysed by both field workers are analysed by the principal investigator who generates his own set of categories and makes his own analyses. The team then meet together; examine the three sets of categories, their applicability and their adequacy; and agree on a common set which they believe are the most appropriate.

The set of categories and a subset of the scripts which have been analysed by the three team members are sent to an external researcher who has no contact with the project or the team. He is not asked to generate a set of categories but to review those submitted by the team in the light of the scripts which have been submitted to him. He meets with the team and reviews the categories and their application and they make whatever revisions are necessary. This is an attempt to ensure that the categories are "internally consistent and mutually exclusive ... leave no apparent logical gaps". (Page 31).

This is not the only procedure and may not be the best but appropriate procedures should be specified. It does not seem to me to be sufficient to say "it is a good idea to have an expert in the area of study verify that the set of categories represents the data adequately and consistently" (Page 31). We need much firmer, much clearer statements about what procedures it is essential to follow if a qualitative evaluation is to be acceptable.

A similar process is necessary at the next stage where conclusions are derived from the data which have thus been analysed. At this stage it is possible to use quantitative techniques. The researcher will have concluded his intuitive sifting, will know what he thinks emerges from his data and will be in a position to test it. If, for example, he has concluded that both parents and young people benefit from the experience of being in a high school class together as now happens, particularly in what we call community schools, he will have the transcripts of all his interviews with both pupils and parents. He will have his categories and he will have his impressions that in general this arrangement is acceptable to both pupils and parents.

> "The involvement of parents in the school for example, may have a number of positive advantages. One pupil described the fact that his mother was doing 'O grade' English as good because she could help him now and that because she knows the school and the teachers she is more likely to understand any problems that might arise. Another pupil whose mother was in the Community School says that her Mum 'knows things she would not know otherwise' and will say things like 'They really care about you at that school': 'if I say something they know what I am talking about. At my last school if you said anything about it, it would not mean anything to them'. (Page 100) (Vallely 1983).

How many pupils and how many parents spoke of this matter in their interviews? How many of them were as positive as the pupils quoted? Were there any negative comments? If so, what were they? How frequently were they made? None of this is in the report that went to the sponsors.

In another project the responses after they have been categorised are summarised numerically. 20% of those interviewed preferred to obtain information by phone but a further 40% preferred highly structured forms of their own. (Dockrell W B and Black H, 1983). A test was done to see whether it was possible to generalise differences found in the sample interviewed to the population. Are teachers, parents, pupils, or wnoever "more likely to" or is it simply that in my sample I have found a difference which I cannot reasonably generalise to the population as a whole. Of course if I had interviewed the whole population then the question of generalisation does not arise.

It is true that in the words of Cronbach "intensive local observation goes beyond discipline to an open-eyed, open-minded appreciation of the surprises natures deposits in the investigative net" (Page 125); However, if the reader and indeedthe researcher himself is to use his head then he must know what the likelihood of a particular response was from a particular group and whether

The pattern of responses observed in a particular sample was likely to be generalisable to a larger sample. It is too easy to assume that one can draw conclusions from an unrepresentative and possibly haphazard set of observations.

Let me emphasise that the standards document does say among the caveats that the evaluator shcuild "recognise the complimentariness of qualitative and quantitative data and interpretation". (Page 132). What I am saying is that this needs to be put in a more positive form together with examples not of how it may be done but of what is the minimum necessary for an acceptable evaluation, one that is not in danger of misinterpretation.

5. The Risk of Collusion

My final point is not a technical one. It is much more general. I might have headed this section "lead us not into temptation" for I refer to the danger of collusion between the evaluator and sponsors. In the standards document there is a guideline which reads, "communicate frequently with representatives of the audiences to ensure that they find the qualitative analyses appropriate and the conclusions and recommendations meaningful". (Page 132). I am tempted to reverse that and say, stay away from your audiences, particularly your paymasters, present or prospective, until you have written your report and then to quote from the same source as my previous quotation, you will be in a position to say "what I have written, I have written". Evaluators in all parts of the world but particularly in some more than others are under pressure, direct or indirect, to produce a report which is acceptable to their audiences and particularly their sponsors. Their next contract or their promotion, their livelihood, or in extreme cases, their lives, are at risk if the report that they produce is not acceptable to their sponsors. It is of course possible to distort quantitative data as British researchers are all too well aware. However, in these days when it is standard practice, if not a formal requirement, that quantitative data be filed with a data archive (as is now the practice in the UK) such manipulation of figures is not as easy as it might have been in the past. I know that it is possible to manipulate computer output but in our field it is rarely worth the effort and ingenuity required. It is now more complex than simply changing correlation figures or mean scores.

It is not yet our practice, though Lawrence Stenhouse (1979) did suggest it, to file qualitative data, field notes, interview typescripts, project diaries or even categorical analyses. It is therefore usually not possible for qualitativ

data to be re-analysed, errors to be discovered and misinterpretations to be corrected.

In these circumstances there is a danger of unintended misinterpretation which derives from pre-dispositions or prejudices or a desire to tone down some findings in order to remain on good terms with audiences or an unconscious aquiescence in the expectations of sponsors. These are subtle and often unconscious pressures. It is easy for us to convince ourselves that we should do things which it is convenient for us to do.

There is an even greater danger in those circumstances where there is a single source of patronage and power, where promotion or even continuation in post is dependent on the goodwill of a single individual or a small group. There the pressures to bend will be even greater than in our own societies. The credibility of qualitative research findings will be nil if they are seen as merely being propaganda. They will have all the credibility of research on smoking sponsored by the cigarette manufacturers. I suggest therefore that what we need are standards, rigorous standards, which cover the gathering, the analysis and the retention of qualitative data, the availability of the data and not merely the reports to a range of audiences, and a statement about how we reconcile all that with the right of the subjects.

6. Conclusion

I have focussed on three issues, none of which are the crucial ones. I have not spoken about the need for perceptiveness and sensitivity in the qualitative researcher. I am aware that the reports of certain qualitative researchers excite me because of their originality and profoundity. I am aware also that there are the reports of other researchers which seem to me to be banal and trivial, to have missed the point completely if there ever was one. That absence of the vital spark seems to me as dangerous to evaluation and to qualitative evaluation as the misleading conclusions which are illustrated in the case quoted in the standards. I do not know how one legislates for perceptiveness or originality.

CHAMBERS P and POWNEY J (1982) "School based curriculum research" British Educational Research Journal, 8,133-140

CRONBACH L (1975) "Beyond the two disciplines of scientific psychology" American Psychology 30, 116-127.

CUMMINGS C, LOWE T, TULIPS J, WAKELING C (1981) Making the Change, Hodder and Stoughton (for SCRE).

DOCKRELL W B and BLACK H (1983) Perceptions of the content and function of school-based assessment End of Grant Report to the Social Science Research Council. Unpublished manuscript, SCRE.

ELLIOT J, BRIDGES I, EBBUT D, GIBSON R, NIAS J (1981) School Accountability Grant McIntyre.

GORDON P (1981) Language and the School Leaver Scottish Council for Research in Education.

HALSEY A H (1972) Educational Priority: E.P.A. Problems and Policies Her Majesty's Stationery Office.

HAMILTON D (1977) In Search of Structure Hodder & Stoughton (for SCRE)

HEWIT G (1982) "A Critique of research methods in the study of reading comprehension" British Educational Research Journal 8, 9-22

JOINT COMMITTEE (1980) Standards for Evaluation of Educational Programs, Projects, and Materials McGraw Hill.

KALLOS D (1980) "On Educational Phenomena and Educational Research" in Dockrell W B and Hamilton D Rethinking Educational Research Hodder & Stoughton.

MacDONALD B and WALKER R (1976) Changing the Curriculum Open Books.

STENHOUSE L (1979) "The Problem of Standards in Illuminative Research", Scottish Educational Review 11 5 - 11.

STRONACH I and WEIR A D (1983) Once Upon a Timetable: Final Evaluation Report on the Clydebank E.E.C. Project. Unpublished Manuscript, Jordanhill College.

VALLELY M and PEACOCK A (1982) Project Report to Lothian Region, Unpublished Manuscript, SCRE.